

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Discord adalah aplikasi media sosial berbasis iOS, Windows, dan Android. Discord merupakan *platform* yang didesain khusus untuk mendukung para pemain *game* untuk dapat berkomunikasi pada saat bermain *game*. Dilansir dari Liputan 6, Discord dirilis pada tahun 2015 oleh Jason Citron dan pada tahun 2020 Discord telah memperoleh rekor pengguna aktif mencapai 100 juta pengguna [1]. Salah satu hal yang menyebabkan Discord dapat digemari oleh penggunanya adalah fiturnya yang dapat mendukung para pemain *game* dalam bekerjasama dalam tim pada saat permainan berlangsung.

Dilansir dari Kompas.com, Kepala Kantor Staf Kepresidenan (KSP) Moeldoko menyatakan media sosial menjadi tempat berbincang masyarakat dan sebanyak 88,4 persen perbincangan tersebut berkaitan dengan suku, agama, ras, dan antar golongan (SARA) [2]. Adanya obrolan yang berkaitan dengan SARA sangat mengganggu kenyamanan *user* dan dapat menimbulkan berbagai permasalahan, mulai dari konflik, kebencian, hingga perpecahan. Dilansir dari Merdeka.com, terdapat kasus pada akun Instagram @rif_opposite yang ditangkap karena menyebarkan ujaran kebencian yang mengandung SARA [3].

Berdasarkan survey yang telah dilakukan melalui Google Form pada 24 April 2022 hingga 26 April 2022 dengan total 55 responden dari usia 20 hingga 30 tahun, didapatkan hasil bahwa 100% responden mengetahui tentang Discord. Pada hasil survey, didapatkan data bahwa 76.4% responden sering menggunakan Discord dalam kegiatan sehari-hari. Kegiatan yang paling sering dilakukan oleh responden menggunakan Discord dengan hasil 83.6% yaitu berkomunikasi dengan *team* saat bermain *game*. Didapatkan data dengan persentase 100% bahwa tipe bahasa yang paling sering digunakan dalam ruang obrolan Discord adalah bahasa informal. Bahasa informal yang sering digunakan dalam Discord dapat memicu timbulnya kalimat berbasis SARA pada ruang obrolan Discord. Hal ini dapat dibuktikan dengan hasil survey yang menyatakan bahwa 63.6% responden menyatakan masih banyak ditemukan kalimat yang mengandung SARA dalam ruang obrolan Discord. Sebanyak 83.6% responden menyatakan bahwa kalimat yang mengandung SARA dalam ruang obrolan Discord dapat mengganggu kenyamanan pengguna dan

dapat memicu dampak negatif dalam obrolan. Menurut responden, beberapa hal yang dapat memicu timbulnya kalimat SARA dalam ruang obrolan discord, yaitu karena kalah dalam bermain *game* sehingga muncul kalimat SARA sebagai luapan emosi, adanya konflik pada saat bermain *game* sehingga memicu timbulnya kalimat SARA, perilaku dan moral pengirim pesan yang terbiasa mengucapkan kalimat yang kurang baik, kurangnya toleransi, serta tingkat pemahaman seseorang akan sebuah permasalahan atau isu sosial yang tergolong rendah.

Berdasarkan permasalahan terhadap banyaknya kalimat SARA pada Discord, maka dibutuhkan penelitian untuk mengidentifikasi apakah kalimat pada ruang obrolan Discord terindikasi sebagai SARA atau bukan. Untuk dapat mengetahui bahwa suatu kalimat dalam obrolan discord merupakan kalimat yang mengandung SARA atau bukan, dibutuhkan adanya suatu sistem klasifikasi data yang dapat mengklasifikasikan kalimat dalam Discord termasuk kalimat SARA atau tidak. Tujuannya adalah untuk membedakan data berdasarkan kelas serta memprediksi kelas dari suatu data.

Penelitian pengklasifikasian data pernah dilakukan sebelumnya dengan menggunakan algoritma K-NN untuk pengklasifikasian spam Bahasa Indonesia [4]. Penelitian tersebut menghasilkan rata-rata *accuracy* sebesar 87.07% dari 4 skenario. Penelitian serupa juga pernah dilakukan dengan algoritma K-NN untuk pengklasifikasian spam pada Instagram dengan tingkat akurasi sebesar 81.67% [5]. Penelitian pengklasifikasian spam pada Instagram juga pernah dilakukan dengan menggunakan Naive Bayes [6]. Dalam penelitian tersebut, menggunakan 1400 *dataset* dan memiliki tingkat akurasi sebesar 80%. Penelitian serupa menggunakan algoritma Naive Bayes juga pernah dilakukan pada sosial media Twitter [7]. Penelitian tersebut berhasil mengklasifikasikan komentar twitter spam dan bukan spam dengan tingkat akurasi yang tinggi sebesar 95.57%. Penelitian dengan menggunakan dua algoritma juga pernah dilakukan, yaitu dengan Support Vector Machine (SVM) dan Naive Bayes Classifier [8]. Kedua algoritma yang digunakan dalam penelitian tersebut menghasilkan tingkat akurasi berkisar antara 70 –79 % dan termasuk dalam kategori baik.

Naive Bayes merupakan algoritma pemrograman sederhana yang menggunakan Teorema Bayes dan efektif digunakan dalam membangun model pembelajaran yang mampu membuat prediksi dengan cepat [9]. Naive Bayes Classifier bekerja sangat baik dibandingkan dengan model classifier lainnya. Hal ini telah dibuktikan melalui jurnal “Naïve Bayes vs. Decision Trees vs. Neural Networks in the Classification of Training Web Pages” mengatakan bahwa Naïve Bayes Classi-

fier memiliki tingkat akurasi yg lebih baik bila dibandingkan dengan model classifier lain [10]. Keunggulan dalam pengklasifikasian data menggunakan Naive Bayes Classifier, yaitu data yang digunakan dapat berupa data kuantitatif maupun kualitatif, penggunaan Naive Bayes dapat menyederhanakan proses desain, pengklasifikasian Naive Bayes tidak membutuhkan *data training* dalam jumlah besar, Naive Bayes dapat dengan cepat membuat keputusan dalam pengklasifikasian data [11].

Berdasarkan keunggulan dari algoritma Naive Bayes Classifier, maka penelitian dilaksanakan menggunakan algoritma Naive Bayes Classifier dalam mengklasifikasikan kalimat SARA pada Discord. Data yang diambil merupakan data dari 5 grup discord dengan total 500 data. Kemudian, data akan dikelompokkan ke dalam dua jenis kategori yaitu SARA atau non SARA. Penyusunan sistem diharapkan dapat mengimplementasikan kalimat SARA pada Discord menggunakan algoritma Naive Bayes Classifier dengan melakukan pengukuran kinerja berdasarkan tingkat *accuracy*, *precision*, *recall*, dan *f1-score*.

1.2 Rumusan Masalah

1. Bagaimana menerapkan algoritma Naive Bayes dalam penelitian pengidentifikasian kalimat SARA pada ruang obrolan Discord?
2. Bagaimana cara mengukur *accuracy*, *precision*, *recall*, dan *F1-Score* algoritma Naive Bayes Classifier dalam pengklasifikasian kalimat SARA pada ruang obrolan Discord?

1.3 Batasan Permasalahan

Penelitian ini memiliki beberapa batasan masalah sebagai berikut:

1. Penelitian ini hanya mendeteksi kalimat SARA yang terdapat pada *Text Channel* di Server Discord, tidak pada *Direct Messages (DM)*.
2. Data yang didapatkan merupakan data yang diambil dengan menggunakan website Export Comments <https://exportcomments.com/> dari 5 grup Discord, terdiri dari grup komunitas terbuka, grup permainan Valorant, grup kelompok belajar, serta grup antarteman sekolah dan kuliah.
3. Kalimat SARA yang dapat dideteksi merupakan kalimat SARA yang dituliskan dalam Bahasa Indonesia, tidak dalam bahasa lain selain Bahasa Indonesia.

4. Klasifikasi kalimat SARA di Discord menggunakan 2 kelas yaitu SARA dan non SARA dengan tipe bahasa formal, sedangkan kalimat dengan tipe bahasa informal akan diubah ke dalam bentuk formal pada tahap *remove slangwords*.

1.4 Tujuan Penelitian

Tujuan dalam penelitian ini, yaitu:

1. Menerapkan Multinomial Naive Bayes dalam mengklasifikasikan kalimat SARA.
2. Menganalisis kinerja algoritma dari tingkat *accuracy*, *precision*, *recall*, dan *F1-Score* berdasarkan hasil *Confusion Matrix*.

1.5 Manfaat Penelitian

Melalui penelitian, diharapkan Discord dapat memanfaatkan hasil penelitian ini untuk mengembangkan sistem filter SARA pada fitur *text channel* dan *Direct Messages (DMs)*.

1.6 Sistematika Penulisan

Berisikan uraian singkat mengenai struktur isi penulisan laporan penelitian, dimulai dari Pendahuluan hingga Simpulan dan Saran.

Sistematika penulisan laporan adalah sebagai berikut:

- Bab 1 PENDAHULUAN
Bab ini menjelaskan mengenai latar belakang penelitian, rumusan masalah, batasan permasalahan, tujuan penelitian, serta manfaat dari penelitian.
- Bab 2 LANDASAN TEORI
Bab ini berisi penjelasan mengenai landasan teori yang digunakan sebagai penunjang penelitian, antara lain pembelajaran mesin, Naive Bayes Classifier, SARA, *Confusion Matrix*, dan Discord.
- Bab 3 METODOLOGI PENELITIAN
Bab ini berisi penjelasan metodologi yang diterapkan dalam penelitian beserta *flowchart* tahapan penelitian, yaitu tahap pengumpulan data (*scrapping*

data), *labeling*, *preprocessing*, *train-test split*, *Feature Extraction*, *classification*, *analyzing model prediction*, dan *evaluation*. Selain itu, terdapat *flowchart* yang menjelaskan alur *preprocessing*, yaitu tahap *case folding*, *data cleaning*, *tokenizing*, *remove slangwords*, *remove stopwords*, dan *stemming*.

- Bab 4 HASIL DAN DISKUSI

Bab ini berisi penjelasan hasil implementasi penelitian. Penjelasan pada bab ini berupa *code* yang digunakan dalam penelitian dan juga *output* yang dihasilkan.

- Bab 5 SIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari penelitian yang telah dilakukan dan saran untuk penelitian berikutnya.

